Subpixel analysis of hyperspectral satellite imagery in case of vegetation classification

Thesis book

Written by

Márton Deák

Supervisor

dr. László Mari, associate professor

Doctoral School

Eötvös Loránd University, Faculty of Science

School of Earth Studies

head: dr. Judit Bartholy D.Sc., professor

Doctoral Program

Geography-meteorology Doctoral Program

head: dr. Dávid Karátson D.Sc., professor

I. Problems regarding the topic

Today more and more hyperspectral data sources are being used, but satellite images have problems than are not relevant in case of aerial or UAV data. Among these maybe I should mention the relatively small spatial resolution (30 m/px in case of Hyperion) compared to the high spectral resolution in the first place. Because of this duality the spectral mixture in every pixel will be high and the ratio of pure data will be small.

The problem of spectral mixture has been analyzed by many scientists who have created different unmixing models, but these aren't mainly usable in case of hyperspectral data because they are made to differentiate between well-separable endmembers in case of multispectral images. Among these is the V-I-S (Vegetation-Impervious-Soil) model to describe urban surfaces – but there are other linear unmixing models mentioned in the thesis.

The other main problem is the high dimensionality which is not a problem only because the analysis requires a lot of processing capacity. Including many bands usually lowers the data pureness and leads to the over-representation of some spectral features. There are some methods to solve this – for example there is the widely used principal component analysis – but they usually underperform in case of hyperspectral data while they also tend to remove some key spectral features needed for vegetation classification.

II. The goals of the research

I had many goals while writing this thesis and all of them aims to solve the problems mentioned above. These are the following:

1. My first goal was to unearth the problems of subpixel classification – ultimately going through the mechanics of spectral mixture models in case of hyperspectral satellite data and vegetation classes. Connecting to this I elaborated on the linear and non-linear unmixing methods.
2. Following this my second goal was to create a method to lower the dimensionality of hyperspectral satellite images while in the meantime the separability of the training areas needed to be higher than before.

3. Building on the abovementioned goals and methods my third goal was to create a spectral unmixing model which works well in case of not well-separable vegetation classes.
4. Using these methods my final goal was to analyze the sample area in the Budai-hills and create a subpixel vegetation map.

III. Methods

My first tool was to create an overview of the already existing methods and go in details regarding their strengths and weaknesses. For this I used the data of the Earth Observing-1 (EO-1) satellite and its hyperspectral camera, the Hyperion. Since I needed classes with low spectral separability to fulfill my goals I chose an image captured on 14. July 2013 where every vegetation class was at the same state in their vegetation cycle – thus the spectral differences caused by sprouts and flowers during the growing phase or different crown-types because of defoliation could be minimized.

The analysis was supplemented by an on-field data collection. For this I chose the János-hill – Csillebérc – Farkas-dike are in the Budai-hills as a sample area. There were different tree species present where the most notable are the black pine (Pinus nigra), beech (Fagus sylvatica), sessile oak (Quercus petraea) and two maple species, the sycamore maple (Acer pseudoplatanus) and field maple (Acer campestre). The vegetation polygons were created by a Trimble Nomad GPS.

For data analysis I used mostly the ENVI 5.3 software but other data processing software (e.g. SPSS) were implemented as well. The calculated spectral distances and classification probabilities were mostly applied in ENVI but they were implemented in other cases – for example by creating the IBA – as well.

The classification was done using the Support Vector Machine (SVM) algorithm where I accepted the classification accuracy as a reliability-index of any analyzed method.

IV. Theses

1. The vegetation patches connected to tree species are appearing as heterogeneous pixels on hyperspectral satellite images, thus they can be unmixed.

The abovementioned statement can be interpreted as the most important foundation of the thesis since it wasn't described and supported anywhere just by the results of this work – but almost everything is based on it. It's important to underline that in this case a pixel counts a homogeneous if the spectral signals it consists from cannot be differentiated by the applied sensor or methodology. Another statement is based on this: heterogeneous pixels are possible to be unmixed while homogeneous ones are not.

According to this every class where it is not possible to tell the spectral signals (and their connecting surface cover types) apart should be merged in the same class. A typical example is a pixel consisting of the abovementioned tree species analyzed by a multispectral sensor which cannot tell the difference between them with high confidence – but it's basically the same if the an area is being captured by a hyperspectral senor but it's dimensions are reduced by principal component analysis (which removes the small differences between the spectral signs) effectively creating a homogeneous pixel from a heterogeneous one.

To my statement it is possible to unmix the data of hyperspectral satellite images – especially the EO-1 Hyperion – but only by using the correct pre-processing methods. These methods and their results are described in the thesis in detail.

2. The subpixel classification of tree species is not only affected by chlorophyll and H_2O but other spectral signals as well.

It's almost obvious that vegetation can be differentiated by its chlorophyll from other surface cover types – and two vegetation types can be told apart by their chlorophyll and H_2O content and their variability through time.

But in case of hyperspectral pixels it is natural that not only one or two attributes are mixing but several other vegetation related or not related ones as well – as described in the thesis. These can be – for example – the lower level vegetation below the tree crowns (in the sample area are these are bushes and other non-arboreal plants) or soil but there are also other non-photosynthetic vegetation parts like the visible parts of branches. In the methods proposed later I described the wavelengths which have a high impact in the classification and their connecting physical-chemical attributes and the evidence shows that many non-photosynthetic elements are affecting the classification as well – which can be caused by the physiognomic variations of the different tree species.

3. I created a band reduction method which lowers the data dimensionality and rises the training spectra separability at the same time (Influential Band Analysis – IBA).

As an answer to the problems presented in the thesis I created a method which takes not only the full image as an input but the training spectra as well. The property of this band selection method is that it selects the bands by considering the separability of the training spectra and raising as much as possible thus also raising the classification accuracy of these classes higher than if the selection was made using the whole image.

I called this method Influential Band Analysis and classified the image using only these bands and compared it to the classification results using the full, the principal component analysis and stepwise discriminant analysis reduced dataset. The results were compared through the classification accuracy and by computing Jeffries-Matusita-distances between them giving in every case better results than in case of the full dataset or other dimensionality-reduced ones. Another positive side of the IBA is that the bands are not dimensionless (like in case of principal component analysis) thus they can be associated with distinct spectral attributes.

4. The V-I-S model used by urban areas can be applied after modifications in case of hyperspectral satellite images used for vegetation analysis.

For spectral mixture analysis I used the V-I-S model as inspiration and modified it so it can be used in case of vegetation endmembers. For this – contrary to the original method – not simple distance (minimum distance) was calculated but I analyzed the SVM and applied its rule images (containing probability values) and used it to create mixed classes – thus it became a non-linear unmixing method.

Using the abovementioned methods I created a decision tree algorithm aiming to use the probability values created by the SVM and created mixed classes. Every pixel was classified using the endmembers to create a mixed class but still kept the opportunity to remain a single-class pixel if only one dominant vegetation species was present.

5. It is possible to classify the pixels showing the Budai-hills on a subpixel level, thus the vegetation map of the area could be created.

Finally, according to the first thesis I examined the methods abovementioned methods to find out what kind of vegetation composition can be even classified using them. The study area selected in the Budai-hills turned out to be an ideal site for this analysis because the natural and artificial effects have created extremely heterogeneous forest patches (both regarding the species composition and the species ratio), which resulted in very diverse spectral signs.

Using the IBA dimension reduction and the spectral unmixing methods I created the subpixel vegetation map of the area, presenting a base for future delineation of forest segments and ecological habitats.

V. Results, conclusions

The main result of the thesis is the produced multistep subpixel classification method. The first step is the dimension reduction process called IBA while the second one is the decision tree algorithm based on SVM aiming to create mixed classes. I stated that after the correct pre-processing steps it is possible to extract the surplus information hidden in the hyperspectral data, thus the heterogeneous (in other cases basically edge) pixels can be now correctly evaluated raising the classification accuracy significantly.

Using this methodology on the sample area and creating classification images the results are clear: classification using the full dataset and classifying for the dominant species yielded an accuracy of 76.67% while implementing the IBA and introducing mixed classes raised the accuracy to 85.78%. I concluded that the vegetation-including pixels of the Hyperion are possible to be unmixed and classified into mixed classes – but this requires the correct pre-processing steps and band selection methods.

VI. Publications

VI.1. Publications used for the base of the theses

Deák M.; Telbisz T.; Árvai M.; Mari L.; Horváth F.; Kohán B.; Szabó O.; Kovács J. (2017). Heterogeneous forest classification by creating mixed vegetation classes using EO-1 Hyperion. *Int. J. Remote Sens. 38*, pp. 5215–5231. https://doi.org/10.1080/01431161.2017.1325529.

Deák M.; Telbisz T.; Árvai M.; Mari L.; Horváth F (2014). Subpixel Vegetation Classification of EO-1 Hyperion Data Through Spectral Reduction, A Case Study in Hungary. In: He, Yingbin; Gao, Mingjie; Zhou, Zhenya (szerk.) *The Application of Remote Sensing and GIS Technology in Crop Production*. Beijing, Kína: China Agricultural Science and Technology Press (CASTP), (2014) pp. 83-88.

Takács K.; Árvai M.; Koós S.; **Deák M.**; Bakacsi Zs.; László P.; Pásztor L. (2017). Sub-parcel terroir mapping supported by UAV-based hyperspectral imagery. *Geophys. Res. Abstr. 19*, Paper: EGU2017-3564 (2017)

Szabó L.; **Deák M.**; Szabó Sz. (2016): Comparative analysis of Landsat TM, ETM+, OLI and EO-1 ALI satellite images at the Tisza-tó area, Hungary, *Acta Geogr. Debr. Lands. and Env. 10:2*, pp. 53-62

Árvai M.; Mészáros J.; **Deák M.**; Koós S.; Takács K.; Szatmári J.; Tobak Z.; Papp L.; Pásztor L. (2018). Természetközeli felszínek hiperspektrális vizsgálata UAV platformról In: Molnár, Vanda (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában IX. : Theory meets practice in GIS, Debrecen, Magyarország : Debreceni Egyetemi Kiadó, (2018) pp. 27-28.

Tulézi F.; **Deák M.**; Kiss K. (2014). Hiperspektrális űrfelvétel alkalmazása talaj szervesanyag-tartalom meghatározására In: Balázs, Boglárka (szerk.) *Az elmélet és a gyakorlat találkozása a térinformatikában V.: Térinformatikai konferencia és* szakkiállítás, Debrecen, Magyarország: Debreceni Egyetemi Kiadó, (2014) pp. 387-395.

Deák, M.; Telbisz T.; Árvai M.; Mari L. (2013). Subpixel vegetation analysis of EO-1 Hyperion satellite images. In: *35th International Symposium on Remote Sensing of Environment*, Paper: 962

Deák M. (2013). The colorful hyperspectrum In: Bircsák, Eszter (szerk.) *Data is Beautiful*: *Data from the view of Society, Science, Art, Design and Technology.* Budapest, Magyarország: Kitchen Budapest (KIBU), (2013) pp. 26-31.

Deák M.; Mari L. (2012). Hiperspektrális űrfelvételezésen alapuló felszínborítási osztályozás problémái a Kárpát-medencében. In: Blanka, Viktória (szerk.) *Kockázat – Konfliktus – Kihívás: A VI. Magyar Földrajzi Konferencia, a MERIEXWA nyitókonferencia és a Geográfus Doktoranduszok Országos Konferenciájának absztrakt kötete.* Szeged, Magyarország: SZTE TTIK Természeti Földrajzi és Geoinformatikai Tanszék, (2012) p. 33 Paper: 33

Deák M.; Ónodi Zs.; Mari L. (2012). Magyarországi falvak lehatárolása EO-1 - hiperspektrális űrfelvételek segítségével. In: Lóki, József (szerk.) *Az elmélet és gyakorlat találkozása a térinformatikában III. - Térinformatikai konferencia és szakkiállítás.* Debrecen, Magyarország: Debreceni Egyetemi Kiadó, (2012) pp. 119-126.

VI.1. Other publications in the scientific field

Deák M.; Kari Sz.; Vizi J.Cs.; Zagorácz M.; Sik A.; Riedel M.; Kohán B.(2018). Examples from the Boundaries of Geographic Survey: Architecture and Flood Modeling. In: Casagrande, G.; Sik A.; Szabó G. (szerk.) *Small Flying Drones Zürich*, Svájc: Springer International Publishing, (2018) pp. 127-156.

Deák M.; Zagorácz M.; Kari Sz.; Sik A.; Riedel M. M. (2017). Pontfelhők az építésügyben – problémák és megoldások. In: Balázs, B (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában *VIII. Theory meets practice in GIS*, Debrecen, Magyarország: Debreceni Egyetemi Kiadó, (2017) pp. 89-96.

Várnagy Sz.; **Deák M.**; Harmath T.; Kari Sz.; Katona A.; Sik A.; Riedel M.M. (2016). Műemléképületek lézerszkenneres 3D felmérésének módszertana és tapasztalatai. *Metszet: Építészeti újdonságok, szerkezetek, részletek: 5*, pp. 50-57.

Telbisz T.; Látos T.; **Deák M.**; Székely B.; Koma Zs.; Standovár T. (2016). The advantage of lidar digital terrain models in doline morphometry compared to topographic map based datasets – Aggtelek karst (Hungary) as an example. *Acta Carst. 45:1*, pp. 5-18.

Mészáros J.; Árvai M.; Kohán B.; **Deák M.;** Nagy B. (2016). UAV based 3D digital surface model to estimate paleolandscape in high mountainous environment. *Geophys. Res. Abstr. 18,* Paper: EGU2016-15430

Koma Zs.; **Deák M.**; Kovács J.; Székely B.; Kelemen K.; Standovár T. (2016). Investigating the limitations of tree species classification using the Combined Cluster and Discriminant Analysis method for low density ALS data from a dense forest region in Aggtelek (Hungary). *Geophys. Res. Abstr. 18,* Paper: EGU2016-10323

Koma Zs.; **Deák M.**; Kovács J.; Standovár T.; Kelemen K.; Székely B.; Kohán B (2016). Alacsony pontsűrűségű LIDAR adatok optimális osztályszámának meghatározása CCDA alkalmazásával erdészeti felmérés példáján. In: Balázs, Boglárka (szerk.) *Az elmélet és a gyakorlat találkozása a térinformatikában VII. Theory meets practice in GIS,* Debrecen, Magyarország: Debreceni Egyetemi Kiadó, (2016) pp. 237-243.

Kohán B.; **Deák M.**; Mészáros J.; Nagy B.; Szabó J.; Szalai Z (2016). Ártéri domborzatfelmérés és elöntésszimuláció: gyorsan, hatékonyan és pontosan – a helyi árvízi védekezés érdekében In: Balázs, Boglárka (szerk.) *Az elmélet és a gyakorlat találkozása a térinformatikában VII. Theory meets practice in GIS,* Debrecen, Magyarország: Debreceni Egyetemi Kiadó, (2016) pp. 229-235.

Székely B.; Telbisz T.; Koma Zs.; Kelemen K.; Szmorad F.; **Deák M.**; Látos T.; Standovár T (2015). Geomorphometric correlations of vegetation cover properties and topographic karst features based on high-resolution LiDAR DTM of Aggtelek Karst, NE Hungary. *Geophys. Res. Abstr.* 17 Paper: EGU2015-12888

Látos T.; Telbisz T.; **Deák M.**; Székely B.; Koma Zs.; Standovár T (2015). Derivation of Doline Morphometrical Data from LiDAR and Contour-based DTMs. *Aragonit 20*: 1 p. 60.

Kohán B.; Timár G.; **Deák M.** (2014). A review of the SRTM digital elevation model and its application in GPS technology. *Tradecraft Review: Periodical of the Scientific Board of the Military Security Office (2012-2014),* pp. 26-37. Paper: 2063-2908

Padányi-Gulyás G.; Stibrányi M.; Mesterházy G.; **Deák M.** (2013). Familiar Road, Unfamiliar Ground. Archaeological Predictive Modelling in Hungary. In: Graeme, Earl et al. (szerk.) *Archaeology in the digital era Vol. II* Amsterdam, Hollandia: Amsterdam University Press

Látos T.; **Deák M.** (2012). Lunar surface analogy simulation of the proposed landing sites of Team Puli Space - Hungarian GLXP Team. In: *European Lunar Symposium 2012 Programme and Abstract Volume* Berlin, Németország (2012) pp. 27-28. Paper: 18.

Látos T.; **Deák M**.; Bérczi Sz. (2012). Landing site modelling for the Puli / Hunveyor-15 lunar rover prototype. In: *43rd Lunar and Planetary Science Conference.* Houston (TX), Amerikai Egyesült Államok : Lunar and Planetary Institute, (2012) Paper: 1748

Deák M. (2011). Landing site analysis for low-budget lunar missions – Landing site candidates of Team Puli Space, participant of the Google Lunar X Prize. In: *42nd Lunar And Planetary Science Conference (LPSC 2011)*, Houston (TX), Amerikai Egyesült Államok : Lunar and Planetary Institute, (2011)

Deák M. (2010). The methodology of finding lava tubes with the use of remote detection on Mars, on the example of a newly found cave. In: *41st Lunar and Planetary Science Conference*. Houston (TX), Amerikai Egyesült Államok : Lunar and Planetary Institute, (2010) Paper: 1507